

Pricing Green Financial Products

Awdesch Melzer*
Wolfgang K. Härdle*
Brenda López Cabrera*

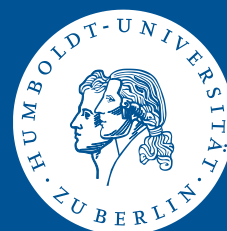


* Humboldt-Universität zu Berlin, Berlin, CEPR, IZA

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



Pricing Green Financial Products *

Awdesch Melzer [†], Wolfgang K. Härdle [‡], Brenda López Cabrera[§]

Monday 14th August, 2017

Abstract

With increasing wind power penetration more and more volatile and weather dependent energy is fed into the German electricity system. To manage the risk of windless days and transfer revenue risk from wind turbine owners to investors wind power derivatives were introduced. These insurance-like securities (ILS) allow to hedge the risk of unstable wind power production on exchanges like Nasdaq and European Energy Exchange. These products have been priced before using risk neutral pricing techniques. We present a modern and powerful methodology to model weather derivatives with very skewed underlyings incorporating techniques from extreme event modelling to tune seasonal volatility and compare transformed Gaussian and non-Gaussian CARMA(p, q) models. Our results indicate that the transformed Gaussian CARMA(p, q) model is preferred over the non-Gaussian alternative with Lévy increments. Out-of-sample backtesting results show good performance wrt burn analysis employing smooth Market Price of Risk (MPR) estimates based on NASDAQ weekly and monthly German wind power futures prices and German wind power utilisation as underlying. A seasonal MPR of a smile-shape is observed, with positive values in times of high volatility, e.g. winter months, and negative values, in times of low volatility and production, e.g. in summer months. We conclude that producers pay premiums to insure stable revenue streams, while investors pay premiums when weather risk is high.

Keywords: market price of risk, risk premium, renewable energy, wind power futures, stochastic process, expectile, CARMA, jump, Lévy, transform, logit-normal, extreme.

JEL code:

*Supported by Deutsche Forschungsgemeinschaft via IRTG 1792, Humboldt-Universität zu Berlin.

[†]Humboldt-Universität zu Berlin. Email: awdesch.melzer@hu-berlin.de

[‡]Humboldt-Universität zu Berlin, C.A.S.E.-Center of Applied Statistics and Economics, Unter den Linden 6, 10099 Berlin, Germany. Research fellow in Sim Kee Boon Institute for Financial Economics, Singapore Management University, 90 Stamford Road, 6th Level, School of Economics, Singapore 178903. Research fellow in Wang Yanan Institute for Studies in Economics(WISE), A307,Economics Building, Xiamen University, Xiamen, 361005 China. Research fellow in National Chiao Tung University, No. 1001, Daxue Rd, East District, Hsinchu City, Taiwan 300.

[§]Humboldt-Universität zu Berlin

1 Introduction

With the increase of renewable power penetration into energy systems, unsteady weather, e.g. clouds and lack of wind, became challenges for a stable electricity supply. Hence, owners of renewable power plants, such as wind or solar farms, face the risk of unstable revenue streams due to volatile weather conditions. These uncertainties may be harmful elements of their business, but with fair risk premium they can trigger investment decisions into new green energy projects. Green financial products allow to transfer risk to investors and create more stable revenues. Indeed, such an insurance-like security (ILS) standardised to volume and delivery period, with a pay-off function dependent on the performance of the underlying will as in stock markets reduce volatility (Pérez-González & Yun 2013) and is widely known as weather derivative.

Weather derivatives (WD) are contracts to hedge risk exposures of weather volatility. The payment is contingent on weather related measurements and is applicable to any renewable energy project, such as wave, solar or wind power projects. Wind speed indices tracking the variations in wind speed, that is required to infer on the wind power production of wind farms, were introduced in Summer 2007 by the U.S. Futures Exchange (USFE) in combination with related futures products for Texas and New York. A first methodology on modelling wind futures based on the Nordix index as underlying has been introduced by Benth & Šaltytė Benth (2009) using risk neutral pricing. The products were not traded well and soon the USFE unlisted them. In 2015, the National Association of Securities Dealers Automated Quotations (NASDAQ) and the European Energy Exchange (EEX) announced to list wind power futures (WPFs) as standardised products responding to the increased interest into renewables in Germany as well as to the growing need of volume risk hedging. NASDAQ introduced WPFs for the German market in December 2015. The EEX followed in October 2016. The market for WPFs is still small but growing and of increasing interest not only to wind power owners, but to banks, investors and owners of conventional power plants (Gersema & Wozabal 2017).

In this article we focus on WPFs traded at the NASDAQ, since the given trading period is larger than we could have collected for the WPFs traded at the EEX. By definition, WPFs are contracts settling against the expected average power production of a future delivery period. The underlying of these contracts at the NASDAQ is the NAREX-WIDE index, a daily average utilisation factor, that translates percent of utilisation directly into euro per wind production hour. Hence, WPFs are linked to WDs not directly via meteorological measurements but rather by measurements affected by meteorological variables.

The field of WD pricing comprises several streams of literature, ranging from econometric pricing

(Campbell & Diebold 2005), equilibrium pricing models (Gersema & Wozabal 2017), burn analysis (Jewson & Brix 2005), indifference pricing (Benth, Ebbeler & Kiesel 2014) to risk neutral pricing (Benth et al. (2007), Benth & Šaltytė Benth (2009), Härdle & López-Cabrera (2012), Härdle et al. (2016), Groll et al. (2016)). Our article contributes to the latter area of research.

Benth et al. (2007) introduced a four step algorithm calculating WD prices for temperature. The methodology accounts for seasonality, a seasonal variance component and employs normalised risk factors by an autoregressive model in continuous time. Benth et al. (2011), Härdle & López-Cabrera (2012) and Groll et al. (2016) show applications in weather markets. Härdle et al. (2016) develop a methodology for spatially and temporally local pricing of temperature dynamics that refines the Benth technology. Methods of modelling rain products were introduced in Härdle & Osipenko (2017) and López-Cabrera et al. (2013). Benth & Šaltytė Benth (2009) make use of a Box-Cox transformation to get CAR models with Gaussian increments or using wavelets (Alexandridis & Zapranis 2013) for the (theoretical) pricing of wind futures and empirical modelling of the underlying.

The underlying wind power utilisation is a complex process. Wind turbines require a certain wind speed to start production and are shut down and disconnected for high wind. These cut-in and cut-off areas can vary across turbine types, but lie usually in the ranges of $3 - 5 \text{ m/s}$ and $25 - 30 \text{ m/s}$, respectively. Due to cut-in and cut-off periods depending on the wind speed, the wind power utilisation exhibits a bi-modal heavily right-skewed density either due to jumps in production and/or strong autocorrelation. The density reduces to a right-skewed distribution when focussing on daily averages.


Recent literature focusses on models with jump diffusion, comprising the heavy tailed risk factors (Barndorff-Nielsen et al. (2013), Weron (2008), Benth & Ortiz-Latorre (2014), Veraart (2016), Schlemm & Stelzer (2012), Benth, Klüppelberg, Müller & Vos (2014)). Jumps in power generation might point into the direction of Lévy-models. If, however, autocorrelation is the main driver of the right-skewed distribution the jump diffusion model will not be a good choice. To tackle this issue, we propose a rather simple framework that allows for generalisation to intraday-modelling and deliver a comparison of our Gaussian candidate model with non-Gaussian alternatives employing information criteria.

Moreover, our candidate model is applicable to other areas of WDs and is able to mimic the inherent stochastic process of the increments in a simple manner. The distribution of wind power is heavily right skewed or shows bi-modality, and hence, a challenge for off the shelf pricing algorithms. Power type transformations, e.g. Box-Cox, help to get closer to normally distributed risk factors (Benth & Šaltytė Benth 2009). Instead of a Box-Cox transformaiton we apply a logit-transformation

to the power utilisation for its better numerical properties. We use periodic B-splines to have a flexible seasonal component, allowing to capture within-year seasonality, that is easily extendible to intra-day estimates (Ziel et al. 2016). In addition, we propose a smooth inter expectile range (sIER) and smooth inter quartile range (sIQR) estimator of the seasonal variance to achieve Gaussian risk factors.

We perform our empirical analysis on a data set consisting of a collection and aggregation of transmission system operator (TSO) data on wind power production and capacity in Germany for each bidding area. We procure quasi maximum likelihood (QML) for the estimation of the CARMA(p, q)-Lévy model coefficients, and an approximation for the Gaussian CAR(p). The Lévy increments show a good fit to the right skewed empirical density of the wind power index indicating increments driven by an asymmetric variance gamma process, leaving however some unexplained autocorrelation in the residuals. Moreover, tails of the distribution are not precisely explained by the variance gamma fit, while the transformed Gaussian model is able to mimic the tails more solid.

In-sample-fit comparison of the models with Akaike information criterion (AIC) for transformed data (Akaike 1978) indicates a preference of the transformed Gaussian over the Lévy type model. The out-of-sample performance of our model for market price prediction deploying root mean squared error (RMSE) and mean absolute percentage error (MAPE). The comparison consists of different market price of risk estimates vs. the burn-in analysis by Jewson & Brix (2005), as well as actual average power utilisation as naive benchmark vs different MPRs. Employing time-variant smooth MPR estimates outperforms the benchmarks: burn analysis and naive actual average utilisation. The data on WPF prices has been collected via Bloomberg Professional for monthly and weekly contracts for 2016.

The rest of the article is structured as follows. First, we briefly discuss wind power indices and summarize the data in Section 2. An introduction to wind power derivatives follows in Section 3. We discuss our empirical findings in Section 4 and conclude with some future research ideas in Section 5. All computations have been done in R and accessible via GitHub and www.quantlet.de.  PGFP

2 Wind Power Index

Wind power indices are the underlyings of WDs and have been in the focus of recent energy research. Ritter et al. (2015b) and Ritter et al. (2015a) calculate a wind power index based on the logarithmic law and a five point non-linear conversion function for hourly NASA MERRA wind speed data. On the application side, Deutsche Börse and EuroWind estimate a wind power index based on MERRA data

and a grid oriented average model for wind power conversion. NASDAQ has introduced the NAREX-WIDE index and Wind Power Futures for Germany in September and December 2015, respectively. Their underlying is calculated in cooperation with the Meteogroup based on ECMWF ERA and NASA MERRA data. The European Energy Exchange AG (EEX) in cooperation with EuroWind has introduced a wind power index based on Deutscher Wetter Dienst (DWD) and Zentralanstalt für Meteorologie und Geodynamik (ZAMG) measurements at a daily resolution. All these indices, except for Ritter et al. (2015b) and Ritter et al. (2015a), are national aggregates.

Most importantly, no reliable index data source for location specific wind power loads is available at a good resolution that provides reasonable estimates. That might be one reason, why the market around WPF is still small, as proximity of the wind turbine to the measurement location of the underlying is essential for reliable hedging.

We calculate the wind power utilisation based on realised production values and installed capacity at a 15min resolution for Germany and a time horizon from 2010 to 2016 published by the transmission system operators (TSO). Wind power production has been monitored in Germany since 2006, however, except for Tennet and 50Hertz, the data has inconsistencies, due to legal untangling of network operation and production in 2010. That restricts our focus to the period starting in 2010 and ending in 2016. We aggregate the production over all four network operator areas and scale by the time-variant capacity factor to get the realised utilisation.

$$U_t \stackrel{\text{def}}{=} \frac{WPL_t}{C_t} \cdot 100, \quad t = 1, \dots, T,$$

where U_t is utilisation in percent and is the ratio of WPL_t the generated wind power load and C_t the capacity of the turbines at time t . Wind power futures are priced on a daily basis, thus we will consider only daily averages of the index. The realised utilisation will be averaged over 96 quarter-hours. Moreover, for simplicity, we remove additional days from leap years, such that we have $T = 2190$ observations. The empirical density function is given in Figure 1.

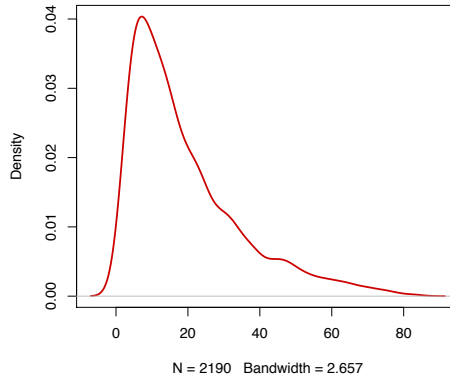


Figure 1: Density of wind power utilisation. Left: Daily TSO data (2010-2015)

3 Wind power derivatives

By definition from the NASDAQ, a wind power future is a contract on settling the expected average wind power production of future delivery periods with the underlying long term utilisation (wind load factor) per contract period. This contract offers the payoff of the value of the underlying asset at the end of the agreement. Similar to temperature derivatives, the contract can be traded at any time until maturity, also within the delivery period $[\tau_1, \tau_2]$ and in comparison to forwards an exchange traded standardised product with daily settlement.

The underlying asset for wind power futures can be calculated as the average wind utilisation from time point τ_1 to τ_2 , the NAREX-WIDE index,

$$NAREX(\tau_1, \tau_2) \stackrel{\text{def}}{=} \frac{100\%}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} \frac{WPL(s)}{C(s)} ds = \frac{100\%}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} U(s) ds, \quad (1)$$

where $WPL(s)$ is the long term wind power load and $C(s)$ is the capacity of the installed power plants and $U(s)$ is the utilisation at time s for a period τ_1 to τ_2 . The percentage of utilisation is directly translated into euro per wind production hour. The WPF is then defined as

$$F_{NAREX}(t, \tau_1, \tau_2) \stackrel{\text{def}}{=} \mathbb{E}^{Q^\theta} \left[\frac{100\%}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} U_s ds | \mathcal{F}_t \right], \quad (2)$$

where t is the current date, τ_1 and τ_2 are the bounds of the delivery period, and $\mathbb{E}^{Q^\theta}[\cdot | \mathcal{F}_t]$ the risk neutral expectation given an information set \mathcal{F}_t comprising all information up to time t .

Let us give a simple example for a WPF. Suppose one is interested in a futures contract for January 2018. The short side of the contract (seller) agrees to deliver electrical wind power generated between January 1st and 31st, theoretically $24 \text{ hours} \times 31 \text{ days} = 744 \text{ MWh}$ for $F_{NAREX}(t, \tau_1, \tau_2) = 20$ euro per MWh and, basically, swaps an uncertain future spot price (direct equivalent average utilisation % within the delivery period) against a certain price (direct equivalent of on average 20 % utilisation within that period). Since the underlying is weather dependent and volatile, the price at time t is calculated using the expected scaled production within that period and is based on historical data and meteorological models that result in the NAREX-WIDE index. In addition to the expected production within the delivery period, the price incorporates all inherent volume risk: if the production is below the expected January-production the long side (buyer) of the contract will make a loss, if the production is above expected January-production the buyer will profit.

3.1 Modelling wind power futures with Gaussian risk factor

A “ready to use” procedure for WDs under risk neutral probability was introduced by Benth et al. (2007) and is presented in Table 1. After decomposing the utilisation U_t into a seasonal component and a stochastic short term variation component, the latter is modeled using an ARMA time series framework as mean-reversion to a market equilibrium. This is followed by the normalisation of the variance component, usually done by a truncated Fourier series as an approximation for the seasonal variance function. Representing the time series as a continuous ARMA model allows for stochastic pricing according to the theory of semi-martingales and gives the market price of risk, the inherent interest rate on the risk exposure.

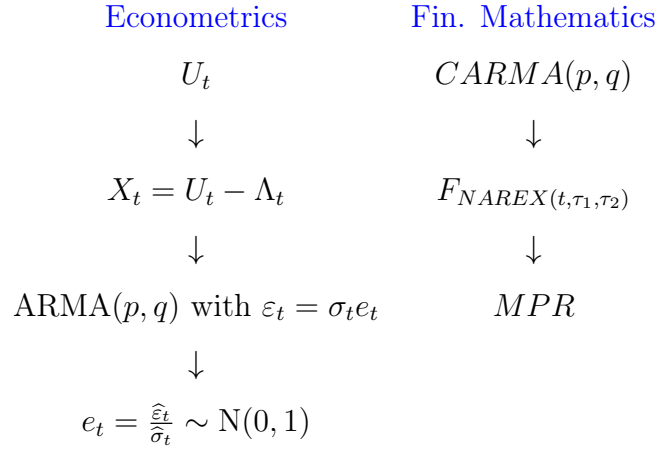


Table 1: Procedure for modelling weather derivatives: deseasonalisation, discrete time series modelling, normalisation of variance component, representation as CARMA model, future pricing and estimation of marginal price of risk.

Modelling WPFs requires more steps than proposed by Benth et al. (2007). To achieve standard Gaussian risk factors, we not only have to deseasonalise and normalise using a seasonal variance estimate, but we need to transform the utilisation process first, since its unconditional density is significantly right skewed. Transformations to normality are logarithmic or power transforms. Logarithmic transformation may however increase the effect from autocorrelated error terms (Benth & Šaltytė Benth 2009). Power and log transformations compress the data around its centre. The power utilisation U_t has a support on $(0, 1)$. By imposing a logit-normal transformation on U_t the new support will be on the interval $(-\infty, \infty)$ (Pinson 2012). The logit-normal transformation and its inverse are given by

$$\begin{aligned}\tilde{U}_t &= \gamma(U) \stackrel{\text{def}}{=} \log \left(\frac{U_t}{1 - U_t} \right) = \Lambda_t + X_t, \quad U_t \in (0, 1), \\ U_t &= \gamma^{-1}(\tilde{U}_t) \stackrel{\text{def}}{=} \{1 + \exp(-\tilde{U}_t)\}^{-1} = [1 + \exp\{-(\Lambda_t + X_t)\}]^{-1}, \quad \tilde{U}_t \in \mathbb{R},\end{aligned}$$

where \tilde{U}_t is the transformed power utilisation with its seasonal mean component Λ_t and short term variation around that mean X_t . Before we discuss how to deal with the transformation of multivariate stochastic processes, we will briefly go through some methods to determine seasonality Λ_t . In literature, truncated Fourier expansions (Alexandridis & Zapranis 2013) and local linear smoothing techniques are common practice, adaptive techniques have been introduced by Härdle et al. (2016).

The deterministic seasonal component can be modelled in several ways. Benth & Šaltytė Benth (2009) and Härdle & López-Cabrera (2012) suggest to use a sinusoidal truncated Fourier series (TFS). Härdle et al. (2016) employ local linear smoothing (LLS) to estimate seasonality and seasonal volatility. Ziel et al. (2016) use periodic B-Splines regression. We model the seasonality and stochastic variance deploying all three methods and compare them wrt their ability to generate Gaussian risk drivers. Generalised Autoregressive Conditional Heteroscedasticity (GARCH(p, q)) is used if required within the TFS set-up. The TFS is defined as

$$\Lambda_t = c_0 + c_1 \cdot t + \sum_{l=1}^L d_l \cos \left\{ \frac{2\pi(t - e_l)}{l \cdot 365} \right\}, \quad (3)$$

$$\sigma_{t,TFS}^2 = c_0 + c_1 \cdot t + \sum_{l=1}^L \left\{ d_{2l} \cos \left(\frac{2l\pi t}{365} \right) + d_{2l+1} \sin \left(\frac{2l\pi t}{365} \right) \right\} + \beta_0 (\sigma_{t-1}^2 \varepsilon_{t-1})^2 + \beta_1 \sigma_{t-1}^2, \quad (4)$$

where Λ_t is the seasonality in dependence of time t , c_0 the intercept, c_1 the linear trend and d_l and e_l seasonality parameters for $l = 1, \dots, L$, β_0 and β_1 are coefficients of the GARCH(p, q) model and ε_t being the residuals from modelling the deseasonalised γ -transformed utilisation factor with an ARMA(p, q) process.

Approximations of the seasonality Λ_t and seasonal variance σ_t^2 with the nonparametric local linear regression are given by

$$\arg \min_{e,f} \sum_{t=1}^{365} \left\{ \bar{U}_t - f_{1s} - f_{2s}(t-s) \right\}^2 K \left(\frac{t-s}{h} \right), \quad (5)$$

and

$$\arg \min_{g,h} \sum_{t=1}^{365} \left\{ \hat{\varepsilon}_t^2 - g_{1s} - g_{2s}(t-s) \right\}^2 K \left(\frac{t-s}{h} \right), \quad (6)$$

respectively. \bar{U}_t is the daily utilisation average over years, f and g are parameters for the local linear regressions, while the kernel $K \left(\frac{t-s}{h} \right)$ evaluates the squared deviation from the local linear regression locally, and $\hat{\varepsilon}_t$ are the residuals from ARMA(p, q)

Moreover, we employ periodic B-splines in capturing seasonality as in Ziel et al. (2016)

$$\Lambda_t = \arg \min_{\alpha_j} \sum_{t=1}^{365} \left\{ \bar{U}_t - \sum_{j=1}^J \delta_j \Psi_j(s_t) \right\}^2, \quad (7)$$

where $\Psi_j(s_t)$ is a vector of known basis functions, δ_j are coefficients, J is the number of knots.

In addition to these methods, we propose a smooth inter expectile range (sIER) as well as a smooth inter quartile range (sIQR) as approximations to the seasonal volatility. Their advantage over standard techniques, is to be flexible in the degree of smoothness, and being able to comprise more variation than the variance does. While the IQR is known to be a robust measure of volatility that comprises 50% of data between the upper and lower quartiles of the distribution irrespective of most outliers (Bowman & Azzalini (1997)), the IER is a measure of the tail variations and coincides with the volatility for $\alpha = \{0, 25, 0.75\}$. Quantile and expectile loss functions go back to Aigner et al. (1976), Newey & Powell (1987) and Breckling & Chambers (1988) who replaced the L_1 norm in the loss function with an L_2 norm. It is defined as

$$\rho_{\alpha,r} = |\alpha - \mathbf{I}\{u < 0\}| |u|^r, \quad r = 1, 2 \quad (8)$$

where for $r = 1$ we have the well known α -quantiles, which are a probability measure at α -level that corresponds to the median for $\alpha = 0.5$, while for $r = 2$ we find the expectile, an α -level moment measure that falls into the mean for $\alpha = 0.5$.

Values beyond $\alpha = 0.5$, however, are somewhat more difficult to interpret for expectiles, whereas any α -quantile can be translated directly into the inverse of the distribution function. To formulate the IER we specify the expectile for an arbitrary location model $Y_t = \theta + \varepsilon_t$ as $e(\alpha, \varepsilon) = \arg \min_{\theta} \mathbb{E}_{\rho_{\alpha,2}}[Y_t - \theta | \mathcal{F}_t]$. The normalised seasonal IER and IQR are defined as

$$\sigma_{IER} \stackrel{\text{def}}{=} \frac{e(\tau = 0.75 | \mathbf{X}) - e(\tau = 0.25 | \mathbf{X})}{2e^{-1}(\tau = 0.75 | \Phi)}, \quad \sigma_{IQR} \stackrel{\text{def}}{=} \frac{\text{med}(|Y - \text{med}(Y)|)}{2\Phi^{-1}(\alpha = 0.75)}.$$

To receive a rather smooth seasonal variance that has the features of creating Gaussian risk factors, we employ smoothing splines with truncated power basis functions, knots at every observation and the roughness penalty approach to tackle on one hand the goodness-of-fit to the data and the smoothness of the fit

$$\arg \min_{\{m, \zeta\} \in \mathbb{R}} \frac{1}{365} \sum_{t=1}^{365} \{\sigma_{t,k} - m(t)\}^2 + \zeta \int dt \left\{ \frac{\partial^2 m(t)}{\partial t^2} \right\}^2, \quad (9)$$

with $k = \{IER, IQR\}$. The first term of (9) addresses the goodness-of-fit while the second term

reduces roughness using the tuning parameter ζ .

Following Wahba (1990) ζ is chosen by cross-validation to minimize the bias

$$\arg \min_{\{m, \zeta\} \in \mathbb{R}} \frac{1}{365} \sum_{\substack{t=1 \\ t \neq s}}^{365} \{\sigma_{t,k} - m(t)\}^2 + \zeta \int dt \left\{ \frac{\partial^2 m(t)}{\partial t^2} \right\}^2. \quad (10)$$

The IER gives us a measure of the conditional volatility at specific α -levels, taking into account the distances of the data above and below the α -level. Since we are not interested in the optimal approximation of the seasonal variance, but rather in an approximation of the variance that generates Gaussian risk drivers, we allow over- or underfitting. To achieve $e_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,sIER}} \sim N(0, 1)$ the objective has to change to minimising the constrained Jarque-Bera test statistic

$$\arg \min_{m, \zeta} \frac{T}{6} \left\{ \left[\frac{T^{-1} \sum_{t=1}^T \left(\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,sIER}} \right)^3}{\left\{ T^{-1} \sum_{t=1}^T \left(\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,sIER}} \right)^2 \right\}^{\frac{3}{2}}} \right]^2 + \frac{1}{4} \left(\left[\frac{T^{-1} \sum_{t=1}^T \left(\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,sIER}} \right)^4}{\left\{ T^{-1} \sum_{t=1}^T \left(\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,sIER}} \right)^2 \right\}^2} \right]^2 - 3 \right)^2 \right\} \quad (11)$$

s.t. smooth seasonal IER or IQR from equation (9).

In a second step we minimize the distance of the standard deviation of the normalised residuals $e_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,sIER}} \sim N(0, 1)$ to the standard deviation of the normal distribution. If normalised residuals were standard normally distributed κ should equal 1

$$\arg \min_{\kappa} \left| \frac{1}{\kappa \cdot \sqrt{T}} (e^\top e)^{\frac{1}{2}} - 1 \right|.$$

After removing the seasonal component Λ_t from the γ -transformed stochastic process \tilde{U}_t , the short term variation is modelled with an Ornstein-Uhlenbeck (CARMA(p, q)) process. Then the multivariate CARMA(p, q) process with $p > q \geq 0$ solves

$$a(D)\mathbf{Y}_t = b(D)D\mathbf{B}(t), \quad D \stackrel{\text{def}}{=} \frac{d}{dt},$$

where the auto-regressive polynomial is given by

$$P(z) = z^p + a_1 z^{p-1} + \dots + a_p$$

and the moving-average polynomial by

$$Q(z) = b_0 + b_1 z^q + \dots + b_{p-1} z^{p-1}.$$

Then we have

$$\tilde{U}_t = \Lambda_t + Y_t \quad (12)$$

$$d\Lambda_t = \Lambda_t dt \quad (13)$$

$$Y_t = \mathbf{b}^\top \mathbf{X}_t \quad (14)$$

$$d\mathbf{X}_t = (\mathbf{A}\mathbf{X}_t + \mathbf{e}_p \sigma_t \theta_t) dt + \mathbf{e}_p \sigma_t dB_t^\theta, \quad (15)$$

where

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ -a_p & -a_{p-1} & \dots & & -a_1 \end{pmatrix}, \quad \mathbf{e}_p = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix}, \quad \mathbf{X}_t = \begin{pmatrix} X_t \\ X_t^{(1)} \\ \vdots \\ X_t^{(p-2)} \\ X_t^{(p-1)} \end{pmatrix},$$

\mathbf{X}_t being the state vector with the derivatives of the continuous-time AR(p) process and $B_t^\theta = B_t - \int_0^T \theta_s ds$ is the Esscher transform to model the risk premium θ using the parametric class of probabilities Q^θ that is equivalent to P (for detail see Karatzas & Shreve (1991), pp 191-201). If $b_j = 0$ for $j \leq 1$ the model reduces to a continuous-time AR(p) model. A generalisation of CARMA processes to $q > p$ are discussed in Brockwell & Hannig (2010).

We get the wind power dynamics as

$$\mathbf{X}_t = \exp\{\mathbf{A}(t-s)\}\mathbf{X}_s + \int_s^t \exp\{\mathbf{A}(t-u)\}e_p \sigma_u \theta_u du + \int_s^t \exp\{\mathbf{A}(t-u)\}e_p \sigma_u dB_u^\theta$$

The mean of the stochastic process is then given by

$$\mu_\theta(s, t, \mathbf{X}_t) \triangleq e_1^\top \exp\{\mathbf{A}(t-s)\}\mathbf{X}_s + \int_s^t e_1^\top \exp\{\mathbf{A}(t-u)\}e_p \sigma_u \theta_u du$$

and the the variance by

$$\Sigma^2(s, t) \triangleq \int_s^t \sigma_u [e_1^\top \exp\{\mathbf{A}(t-u)\}e_p]^2 du.$$

To determine the futures prices of transformed stochastic processes we follow Benth & Šaltytė Benth (2009). Since the futures prices are determined by the average of utilisation percent within the contract period we can write

$$\begin{aligned} \hat{F}_{t, \tau_1, \tau_2} &= \mathbb{E}_\theta [U_t | \mathcal{F}_t] \\ &= \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} \left(1 + \exp \left[- \left\{ \Lambda_t + \mu_\theta(s, t, \mathbf{X}_t) + \frac{1}{2} \Sigma^2(s, t) \right\} \right] \right)^{-1}. \end{aligned}$$

Then the back-transformed power utilisation is given by the inverse transformation of the sum of seasonality, the drift and the diffusion component

$$U_t = (1 + \exp[-\{\Lambda_t + \mu_\theta(s, t, \mathbf{X}_t) + \Sigma(s, t)Z\}])^{-1},$$

where $Z \sim N(0, 1)$ is white noise.

3.2 Modelling wind power futures with jump-risk

To incorporate jump risk we model a \mathbb{R}^m -valued Lévy process \mathbf{L} that satisfies squared integrability $\mathbb{E}[\|\mathbf{L}\|^2] < \infty$. The advantage of the Lévy-type model is to model the inherent jump structure of U_t directly without imposing a transformation as

$$U_t = \Lambda_t + Y_t$$

for which the Ornstein-Uhlenbeck (CARMA(p, q)-Lévy) process can be represented as

$$d\Lambda_t = \Lambda_t dt \tag{16}$$

$$Y_t = \mathbf{b}^\top \mathbf{X}_t \quad \text{state equation} \tag{17}$$

$$d\mathbf{X}_t = (\mathbf{A}\mathbf{X}_t)dt + \mathbf{e}d\mathbf{L}_t, \quad \text{observation equation} \tag{18}$$

where

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots \\ 0 & \dots & \dots & 0 & 1 \\ -a_p & -a_{p-1} & \dots & & -a_1 \end{pmatrix}, \quad \mathbf{e}_p = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix}, \quad \mathbf{X}_t = \begin{pmatrix} X_t \\ X_t^{(1)} \\ \vdots \\ X_t^{(p-2)} \\ X_t^{(p-1)} \end{pmatrix},$$

\mathbf{X}_t being the state vector with the derivatives of the continuous-time AR(p) process. If $b_j = 0$ for $j \leq 1$ the model reduces to a continuous-time AR(p) model. A generalisation of CARMA processes to $q > p$ are discussed in Brockwell & Hannig (2010). \mathbf{L} is the pure jump Lévy process with the Lévy-Itô representation $\mathbf{L}_t = \int_0^t \int_0^\infty z N^L(ds, dz), t \in [0, T]$, where N^L is a Poisson random measure with Lévy measure l that satisfies $\int_0^\infty z l(dz) < \infty$ (Benth & Ortiz-Latorre 2014). If all eigenvalues of the coefficient matrix \mathbf{A} have negative real parts, then $(\mathbf{X}_t)_{t \in \mathbb{R}}$ given by

$$\mathbf{X}_t = \int_{-\infty}^t \exp\{\mathbf{A}(t-s)\} \mathbf{e}_p d\mathbf{L}_s$$

is the strictly stationary solution to the SDE (18) and

$$Y_t = \mathbf{b}^\top \mathbf{X}_t = \int_{-\infty}^t \mathbf{b}^\top \exp\{\mathbf{A}(t-s)\} \mathbf{e}_p d\mathbf{L}_s \quad (19)$$

represents the CARMA(p, q) process without stochastic volatility (Barndorff-Nielsen et al. 2013). In general Lévy processes are semi-stationary processes. From economic rational, however, we rather want to model prices that vary around an equilibrium, hence the mean-reverting part of the CARMA model. For a discussion of L_2 -integrability conditions of Lévy semi-stationary (LSS) processes see Barndorff-Nielsen et al. (2013), for this research we assume integrability conditions to hold true for the presented CARMA(p, q) model.

To model the risk premium, we need to consider the Esscher transform for semi-martingales (Shiryaev (2001), pp. 701-703, Barndorff-Nielsen et al. (2013)), that is defined via a local Q -martingale Q_L^θ for \mathbf{L} and the real valued parameter function $\theta(t)$ through the Radon-Nikodym density process

$$\left. \frac{dQ_L^\theta}{dP} \right|_{\mathcal{F}_t} = \exp\left(\int_0^t \theta(s) d\mathbf{L}_s - \int_0^t \phi_L\{\theta(s)\} ds\right),$$

with

$$\phi_L(x) = \exp[\mathbb{E}\{xL_1\}] = \psi(-ix) = dx + \frac{1}{2}x^2b + \int_{\mathbb{R}} \{\exp(xz) - 1 - xz\mathbf{1}_{\{|z|\leq 1\}}\} l_L(dz).$$

Following Shiryaev (2001) and Barndorff-Nielsen et al. (2013), the Lévy properties are preserved in the Lévy-Khinchine representation for

$$d_\theta = d + b\theta + \int_{|z|\leq 1} z\{\exp(xz) - 1\} l_L(dz),$$

and having the mean adjusted moment generating function

$$\psi_L^\theta(x) = \psi_L(x + \theta) - \psi_L(x).$$

For an arithmetic price model with

$$U_t = \Lambda_t + Y_t$$

and the Esscher transform measure Q^θ we find the future price given by

$$F_t(T) = \Lambda_T + \left[\int_{-\infty}^t b^\top \exp\{A(T-s)\} d\mathbf{L}_s + \mathbb{E}_Q[\mathbf{L}_1] \int_t^T b^\top \exp\{A(T-s)\} \theta(s) ds \right].$$

4 Empirical results

4.1 The Gaussian case

The coefficient estimates of the seasonal component modelled with truncated Fourier series are presented in Table 2. The coefficient estimate \hat{c}_1 in Table 2 is almost 0 indicating stationarity of the underlying wind power utilisation process U_t . Augmented Dickey–Fuller (ADF) and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests provide the same conclusion with p -values below 0.01 and above 0.1, respectively. For the local linear regression we use rule-of-thumb (RoT) bandwidth $\hat{h}_{RoT} = 2.946$ as suggested by Bowman & Azzalini (1997) on page 31, as well as allowing some oversmoothing (RoTos) and bias minimising cross-validation (CV). To extract the seasonality using periodic B-splines we estimate six coefficients, all of which are negative as Table 3 shows.

\hat{c}_0	\hat{c}_1	\hat{d}_1	\hat{d}_2	\hat{e}_1	\hat{e}_2	\hat{h}_{RoT}	\hat{h}_{RoTos}	\hat{h}_{CV}
−1.749	0.000	0.474	9.275	0.083	−17.970	2.946	8.140	0.823

Table 2: Coefficient estimates for the seasonality approximation with truncated Fourier series (TFS) and bandwidth: rule-of-thumb (RoT), oversmooth RoT (RoTos) and cross-validated (CV).

$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\alpha}_6$
−1.685	−2.402	−1.992	−1.760	−0.766	−1.800

Table 3: Coefficient estimates for the seasonality approximation with periodic B-Splines method.

The estimated seasonality curves are presented in the left panel of Figure 2. While the truncated Fourier series and B-splines method seem to give good fit of the seasonal component, local linear smoothing performs well for larger bandwidths than proposed by cross-validation.

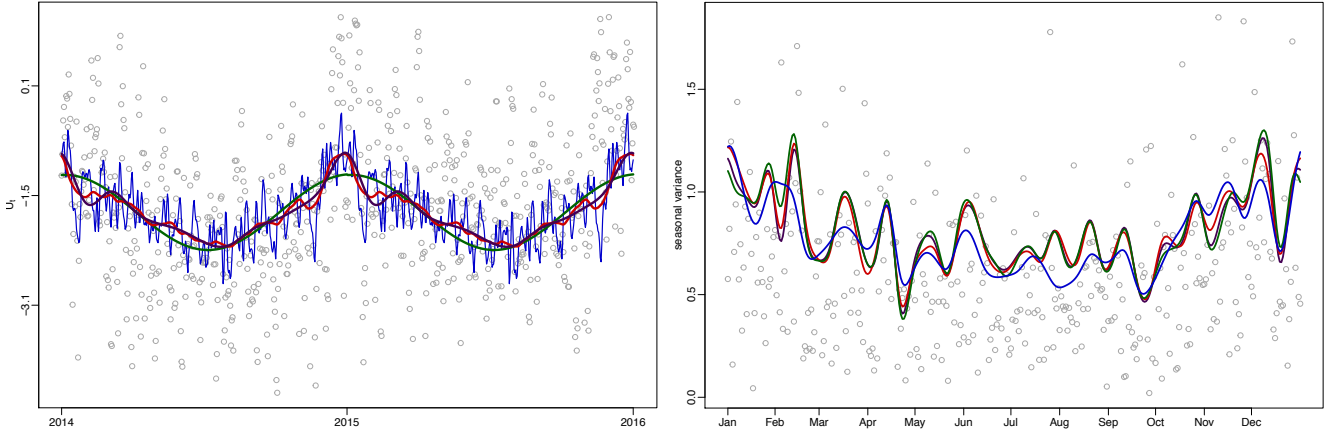


Figure 2: Seasonality and seasonal variance: **Local Linear Estimator with CV-bandwidth**, **LLE-oversmooth**, **Truncated Fourier Series**, **periodic B-Splines**

The seasonal variance is captured by (4) representing the TFS, (6) giving the LLE and 9 for the smooth IER. Estimated curves are presented in the right panel of Figure 2. Coefficient estimates of the TFS are given in Table 4. An ARCH-test indicates no requirement of volatility GARCH-modelling, with p -value of 0.844. The bandwidth for the estimate $\hat{\sigma}_{t,LLE}$ has been chosen via the rule of thumb by Bowman & Azzalini (1997), page 31. Bandwidth selection using cross-validation in the cases of seasonality as well as seasonal variance leads to very coarse and very smooth approximations, respectively. We estimated the IER seasonal variance using cross-validation and maximisation the normality as tuning parameter choices. Maximisation of normality is done in two steps. First, we estimate the tuning parameter ζ that maximises normality with respect to third and fourth moments, then we rescale the seasonal variance such that the normalised increments follow standard normal distribution. We select a scaling parameter κ that minimises the absolute deviation of the standard error of $e_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,sIER}}$ from 1

$$\arg \min_{\kappa} \left| \frac{1}{\kappa \cdot \sqrt{T}} (e^\top e)^{\frac{1}{2}} - 1 \right|.$$

\hat{c}_0	\hat{c}_1	\hat{d}_1	\hat{d}_2	\hat{d}_3	\hat{d}_4
0.591	0.000	0.131	0.013	0.032	-0.033
\hat{h}_{CV}	\hat{h}_{RoT}	\hat{h}_{RoTos}	$\hat{\zeta}_{CV}$	$\hat{\zeta}_{\Phi}$	$\hat{\kappa}$
33.541	2.143	6.428	0.844	0.1775	1.639

Table 4: Top: Coefficient estimates for the approximation of the seasonal variance with truncated Fourier series (TFS). Bottom: tuning parameters for LLE, smooth IQR seasonal variance estimates.

The tuning parameters were $\hat{\zeta}_{CV} = 0.844$ and $\hat{\zeta}_{\Phi} = 0.178$ were used to estimate seasonal variance for

IER and IQR. After rescaling of the residuals with $\hat{\kappa} = 1.639$ we get standard normal increments. The results of the normalisation using different approximations of the seasonal variance are presented in Figure 3. We use Jarque-Bera-test (JBT), Andersen-Darling-test (ADT), Shapiro-Wilk normality test (SWT), Cramer von Mises test (CvM), Kolmogorov-Smirnov test (KST) to measure the performance in achieving Gaussian risk factors. The labels in Figure 3 are to be read as, for instance, in "TFS-IQR.CV": the first term stating the seasonality extraction method, the second part the seasonal variance estimation technique. Seasonality extraction using TFS and seasonal variance normalisation employing IQR with CV tuning parameter. The choice of seasonal variance approximation has large influence on the outcome. The robust smooth IQR outperforms any other method if the tuning

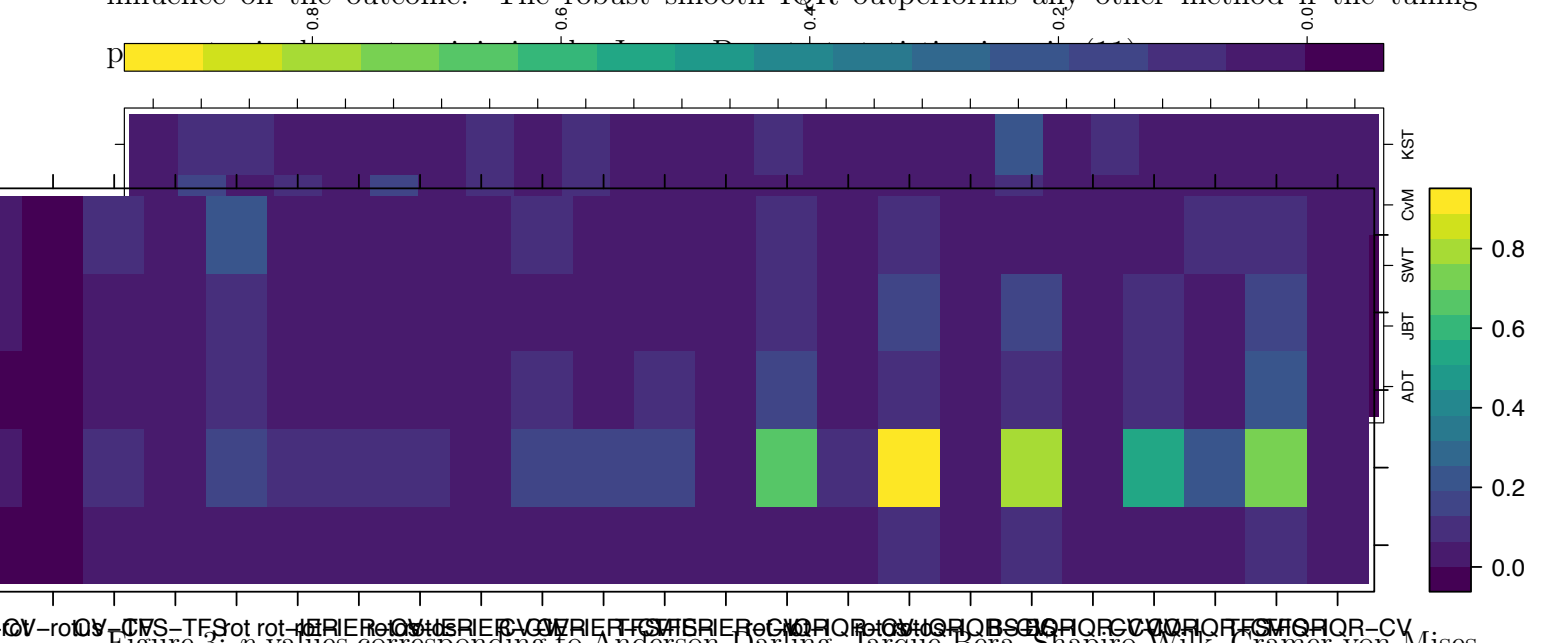


Figure 3: p -values corresponding to Anderson-Darling, Jarque-Bera, Shapiro-Wilk, Cramer-von-Mises and Kolmogorov-Smirnov normality tests for several methods of seasonal variance normalisation.

We model the short-term variation using an $\text{ARMA}(p, q)$ model in the discrete world from which we infer to the continuous time ARMA parameters. While we estimate the coefficients of $\text{CAR}(p)$ models with an approximation, we estimate the $\text{CARMA}(p, q)$ coefficients employing the Quasi Maximum Likelihood Estimation (QMLE). A nice overview on QMLE for CARMA -Lévy processes is provided in Schlemm & Stelzer (2012). The autocorrelation function (ACF) and partial ACF indicates an $\text{ARMA}(4, 2)$ representation (see Figure 4). Following the Bayesian Information Criterion (BIC) that penalises large models suggests an $\text{AR}(3)$ model. The original series in the Lévy-jump representation is modelled then by an $\text{ARMA}(3, 1)$ as the jump process requires $q \geq 1$.

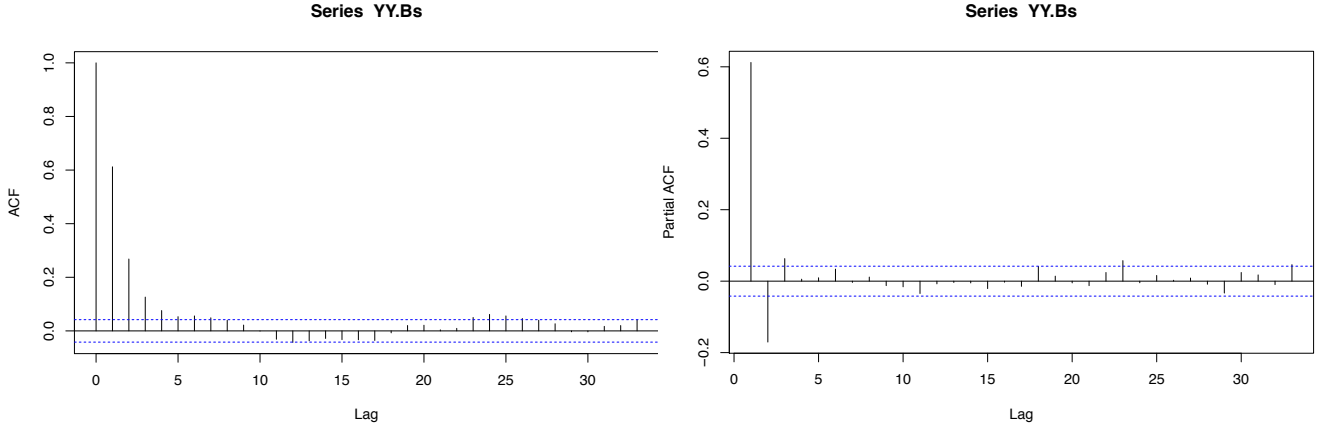


Figure 4: ACF and PACF of the deseasonalised time series

Coefficient estimates for $\text{ARMA}(3, 0)$ are given by $\hat{a}_1 = 0.724$, $\hat{a}_2 = -0.216$ and $\hat{a}_3 = 0.060$. The MA parameter from CARMA-Lévy modelling is $b_1 = -0.551$. Fitting the sIQR-normalised increments to the standard normal distribution gives an Akaike Information Criterion of 6218.001 (This value will be used later in the comparison with the CARMA-Lévy model). Normalising the residuals with the seasonal variance results a very good fit, as the density, log-density and quantile-to-quantile matching in Figure 5 suggests.

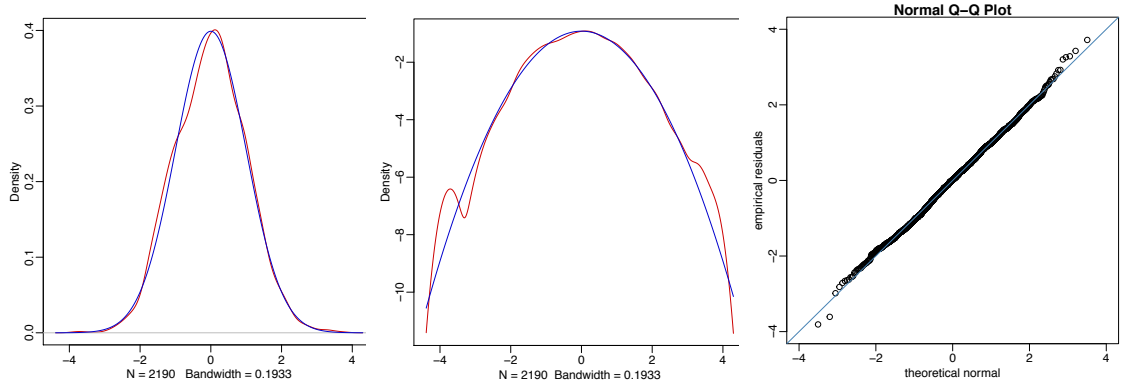


Figure 5: From left to right: pdf of the standardised residuals (with IQR) of the logit transformed utilisation factor vs normal pdf, log-pdf vs. log-normal pdf, quantile-quantile plot.

4.2 The stationary Lévy case

Modelling the untransformed deseasonalised utilisation factors, we consider the generalised hyperbolic (GH) distribution, introduced by Barndorff-Nielsen (1978), to capture the features of the Lévy-process. We follow Barndorff-Nielsen et al. (2013) in their representation. Suppose \mathbf{X} is a k -dimensional random vector, then

$$\mathbf{X} \stackrel{\text{law}}{=} \mu + W\gamma + \sqrt{W}\mathbf{A}\mathbf{Z},$$

where $\mathbf{Z} \sim N_k(\mathbf{0}, \mathbf{I}_k)$, $\mathbf{A} \in \mathbb{R}^{d \times k}$, $\mu, \gamma \in \mathbb{R}^d$. W is a non-negative random variable, following Generalised Inverse Gaussian (GIG) distribution and independent of \mathbf{Z}

$$f_{GIG}(x) = \left(\frac{\psi}{\chi}\right)^{\frac{\lambda}{2}} \frac{x^{\lambda-1}}{2K_\lambda(\sqrt{\chi\psi})} \exp\left\{-\frac{1}{2}\left(\frac{\chi}{x} + \psi x\right)\right\},$$

where K_λ is the modified Bessel function of third kind and index λ

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda-1} \exp\left\{-\frac{x}{2}(y + y^{-1})\right\} dy,$$

while the parametrisation of ψ and χ are given by

$$\psi = \bar{\alpha} \frac{K_{\lambda+1}(\bar{\alpha})}{K_\lambda(\bar{\alpha})}, \quad \chi = \frac{\bar{\alpha}^2}{\psi} = \bar{\alpha} \frac{K_\lambda(\bar{\alpha})}{K_{\lambda+1}(\bar{\alpha})}.$$

For the one dimensional data ($d = k = 1$), we restrict the estimation of parameters to the distributions in Table 5

Distribution	λ	$\bar{\alpha}$	χ	ψ
GHYP	$\lambda \in \mathbb{R}$	$\bar{\alpha} > 0$	$\chi > 0$	$\psi > 0$
NIG	$\lambda = -\frac{1}{2}$	$\bar{\alpha} > 0$	$\chi > 0$	$\psi > 0$
t (ν df)	$\lambda = -\frac{\nu}{2} < 1$	$\bar{\alpha} = 0$	$\chi > 0$	$\psi = 0$
HYP	$\lambda = -\frac{d+1}{2}$	$\bar{\alpha} > 0$	$\chi > 0$	$\psi > 0$
VG	$\lambda > 0$	$\bar{\alpha} = 0$	$\chi = 0$	$\psi > 0$

Table 5: Bounds of parameters for the generalised hyperbolic (GHYP), normal inverse Gaussian (NIG), Student- t with ν degrees of freedom, hyperbolic (HYP) and variance gamma (VG) distributions.

The fitting to the distributions in Table 5 for the symmetric and asymmetric cases as well as for the Gaussian distribution have been done using quasi-maximum likelihood estimation (QMLE).

First we estimate of the CARMA(3, 1)-Lévy process using the QMLE, then we fit the increments to potentially interesting distributions. Results of fitting the CARMA-Lévy increments based on the deseasonalised utilisation factor to the 5 distributions, including symmetric and asymmetric cases, are given in Table 6. Comparing the fits to the selected distributions by the Akaike Information Criterion indicates consideration of the Variance-Gamma distribution to be best-suited to represent the inherent distribution of the Lévy-increments.

To make it possible to compare models of transformed and non-transformed data, Akaike (1978) suggests adjusting the AIC of the transformed model with the Jacobian of the transformed data. In our case it is simply adding

$$\text{AIC}_\gamma + 2 \sum_{t=1}^T \log \left\{ \frac{\partial \log \left(\frac{U_t}{1-U_t} \right)}{\partial U_t} \right\} = \text{AIC}_\gamma + 2 \sum_{t=1}^T \log \left(\frac{1}{U_t - U_t^2} \right),$$

resulting in $\text{AIC}_\gamma + \text{Adjustment} = 6218.001 + 2 \cdot (4847.232) = 15912.47$. Following Akaike (1978), we can directly conclude that the transformed model is preferred over any of the generalised hyperbolic CA

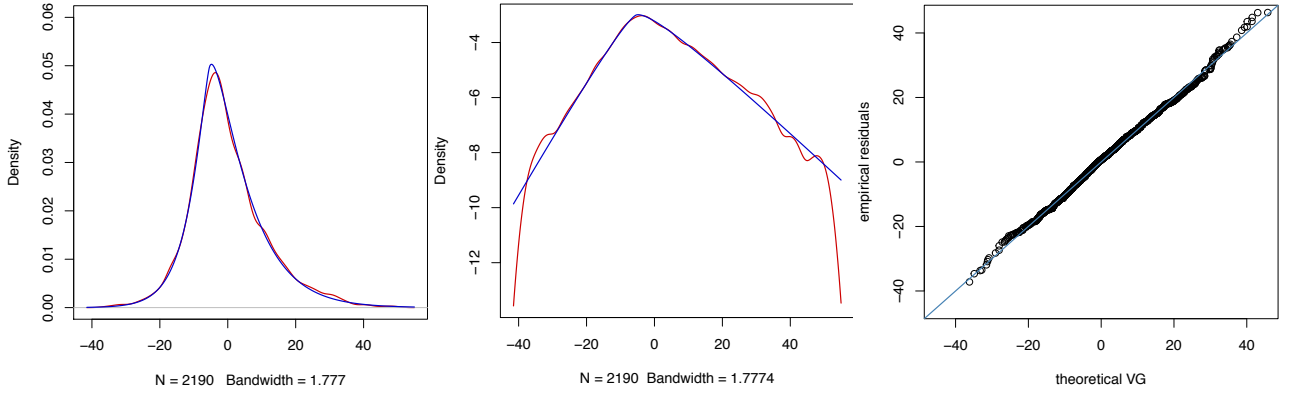


Figure 6: From left to right: pdf and histogram of the standardised residuals vs pdf of a VG distribution, log-pdf vs. log-VG pdf, quantile-quantile plot.

Moreover, the log-density fit in Figure 7 indicates that extreme values of our wind power utilisation are not well described by the Variance-Gamma fit. Further, the the ACF of squared residuals indicates existence of some autocorrelation and thus misspecification in the Lévy case (see Figure 7). Thus, for the estimation of WPF prices, the calibration of the MPR we will use the Gaussian logit-transformed

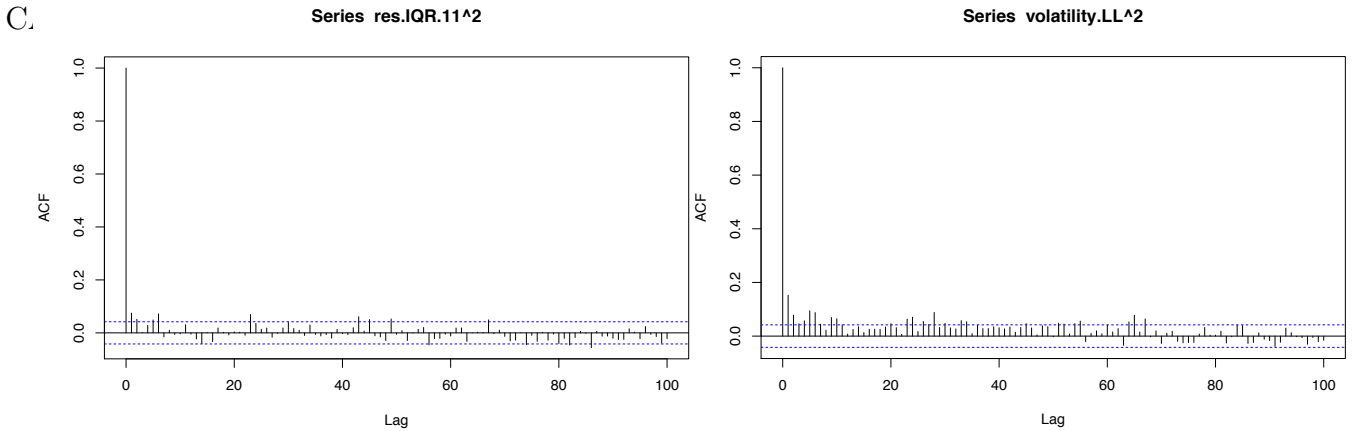


Figure 7: ACF of squared residuals: Gaussian (left), Lévy (right).

Model	Sym	λ	$\bar{\alpha}$	μ	σ	γ	AIC	LLH	Conv	Iter
VG	F	1.426	0.000	-5.386	10.410	5.390	16530.366	-8261.183	T	319
GHYP	F	1.418	0.052	-5.365	10.416	5.367	16532.365	-8261.182	T	404
HYP	F	1.000	0.008	-4.580	10.886	4.587	16547.773	-8269.886	T	167
NIG	F	-0.500	1.229	-5.663	10.308	5.664	16551.142	-8271.571	T	177
t	F	-2.908	0.000	-6.367	10.251	6.468	16580.033	-8286.016	T	205
HYP	T	1.000	0.691	-1.329	11.302	0.000	16667.974	-8330.987	T	92
GHYP	T	0.600	0.859	-1.319	11.331	0.000	16669.597	-8330.799	T	155
NIG	T	-0.500	1.048	-1.239	11.368	0.000	16669.612	-8331.806	T	82
VG	T	1.541	0.000	-1.380	11.268	0.000	16670.515	-8332.257	T	78
t	T	-2.252	0.000	-1.084	11.577	0.000	16680.239	-8337.120	T	88

Table 6: Estimation results: Fit of CARMA-Lévy increments to a selection of generalised hyperbolic distributions.

4.3 Market price of risk

In this section we discuss the market price of risk θ , a parameter that essentially captures the additional return for taking more risk. θ can be inferred from the difference of the historical (under P) and risk-neutral (under Q) behaviour of the underlying (Härdle & López-Cabrera 2012) given liquidity of the derivative market. The wind power futures market is rather young, hence, liquidity cannot be guaranteed. Nevertheless, we assume the same mechanisms to be working at least approximately.

We investigate the MPR structure within contracts over all trading days to understand the effect of time to maturity on the risk premium. From electricity futures it is well known that the risk premium is positive close to delivery, decaying and even becoming negative the further the delivery is in the future (Benth, Klüppelberg, Müller & Vos 2014). Positive risk premiums are associated with the need for locking in prices (here equivalent to volume) as a hedge against weather or spike risk. Similar results can be expected for WPF as their underlying directly translates into price per MWh and is not storable at low cost. Moreover, we are interested in the seasonal structure of WPF, especially, their seasonal behaviour. It is expected that in times of high volatility investors tend to pay a risk premium to hedge against weather or spike risk (Härdle & López-Cabrera 2012). This coincides with the definition of the risk premium

$$RP_{\tau_1^i, \tau_2^i}^i \triangleq \int_{\tau_1^i}^{\tau_2^i} \theta_u \sigma_u \mathbf{e}_1^\top \mathbf{A}^{-1} [\exp\{\mathbf{A}(\tau_2^i - u)\} - \mathbf{I}_p] \mathbf{e}_p du, \quad (20)$$

where θ is directly connected to seasonal volatility σ_t .

To value wind power futures we consider a constant and a time-variant smooth MPR. Since wind power production is strongly affected by current wind conditions, it is intuitive to investigate short term contracts, such as weekly and monthly WPF. The weekly contracts are set over 10 days, while the monthly contracts have a 4 weeks delivery period. We observe the underlying wind power every day. Prices are set only on trading days. This leads to a mismatch resulting in less than 10 or 30 days per contract type. Thus, on average 7 to 8 days of weekly futures capture the information of the 10 days contracts and 20 days of monthly futures capture information of 30 days contracts. Since only little trading happens until the delivery period begins, we focus on the delivery period.

We estimated the empirical market price of risk θ as the volatility-scaled difference between the futures prices at the NASDAQ for monthly and weekly contracts and our expected value over the contract period. The constant MPR is estimated for different futures contracts with different measurement periods $[\tau_1^i, \tau_2^i], t < \tau_1^i < \tau_2^i$, for $i = 1, \dots, I$ contracts

$$\hat{\theta}_t^i = \arg \min_{\theta_t^i} \left(F_{NAREX}(t, \tau_1^i, \tau_2^i) - \hat{F}_{NAREX}(t, \tau_1^i, \tau_2^i) \right)^2.$$

A smooth MPR estimates can be achieved by minimising (21) for given empirical implied MPR estimates $\hat{\theta}_t$.

$$\arg \min_{\{f, \zeta\} \in \mathbb{R}} \sum_{t=1}^n \left\{ \hat{\theta}_t - f(u_t) \right\}^2 + \zeta \int dt \left\{ \frac{\partial^2 f(u_t)}{\partial t^2} \right\}^2 \quad (21)$$

The empirical MPR estimates are presented in Figure 8 for all contracts and trading days. The smoothed median is illustrated in red and the inter quartile range (IQR) as the grey shaded area. The left panel shows $\hat{\theta}_t$ for weekly WPF. The MPR is mostly positive for weekly and monthly contracts. It shows a curve shape, converging to zero at the beginning of the delivery period, which leaves approximately 7-8 trading days for weekly and at 20 trading days montly contracts. Hence we observe a jump in the θ at these bounds.

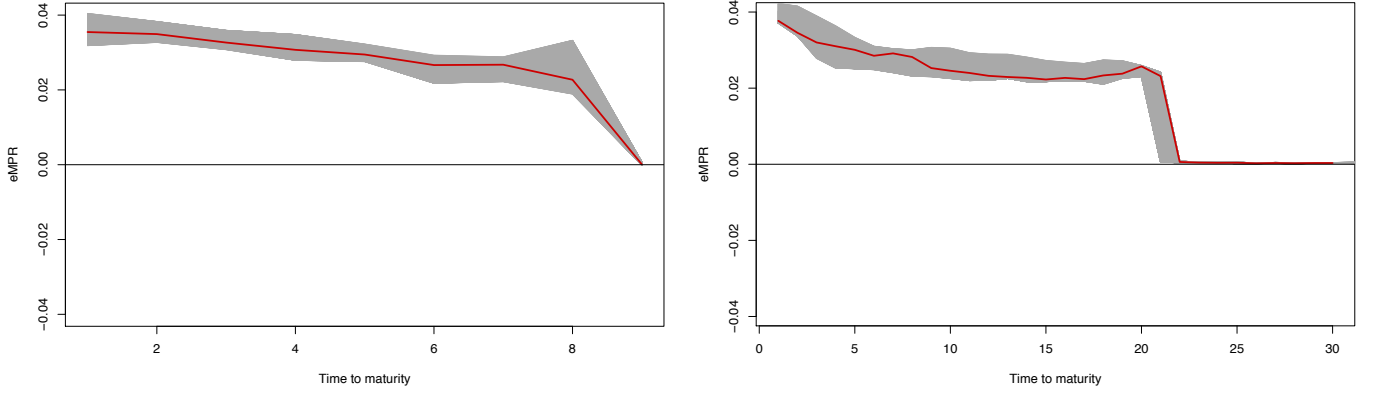


Figure 8: Left: empirical Market Price of Risk (MPR) $\hat{\theta}$ for weekly contracts. Right: $\hat{\theta}$ for monthly contracts. Median $\hat{\theta}$ in red, IQR shaded in grey.

To observe the seasonal structure of WPFs we estimate $\hat{\theta}$ contract-wise, resulting in a smile-shape for weekly contracts. Moreover, we find a change of sign in the summer months, as Figure 9 reveals. This points towards investors paying a premium in the winter months, when seasonal volatility is high, and producers paying a premium in summer months, when volatility but similarly production is low, to insure their production. As for monthly contracts, we have only limited data, we observe an increasing structure, beginning negative and ending in the positive over a period slightly more than half a year. Nevertheless we can assume that the structure will be similar for the winter months January to March, leading to a smile shape. This indicates that the MPR is strongly affected by seasonal structure.

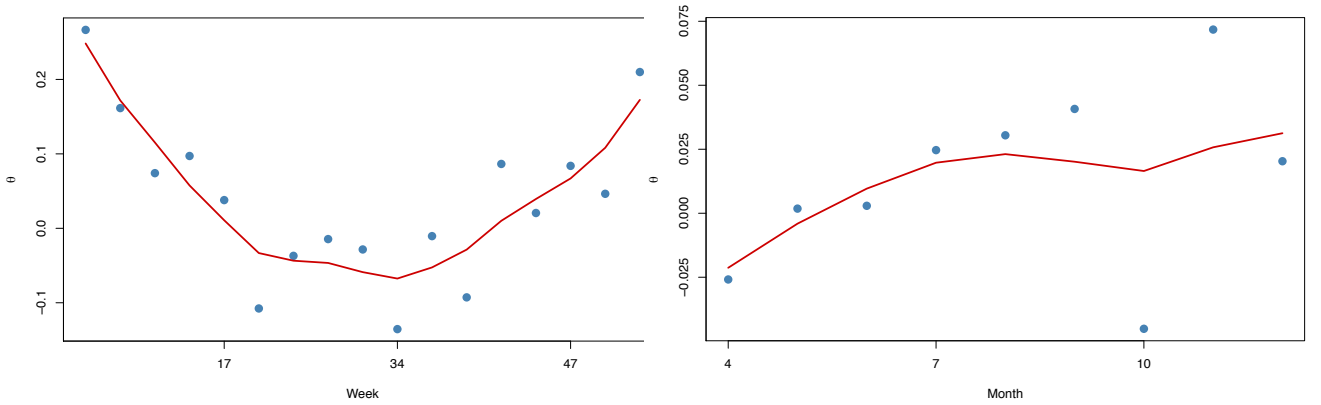


Figure 9: Seasonal structure of MPR. Left: θ for weekly contracts. Right: θ for monthly contracts. Samples consists of 17 weekly and 9 monthly contracts.

In Table 7 we present out-of-sample backtesting results of $\theta = \{0, \text{OLS}, \text{smooth}\}$ employing the performance measures RMSE and MAPE, given in (22) and (23), respectively, for weekly WPF

contracts.

$$RMSE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \sqrt{(y_i - \hat{y}_i)^2} \quad (22)$$

$$MAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (23)$$

When comparing the different estimates of MPR among our models, the smooth MPR performs best, since it captures the temporal dynamic that exists even after normalisation of the risk premium with seasonal variance. Also comparing the futures price estimates with the actual average production within the delivery period, we find slightly better performance of the smooth MPR estimate. We conclude that a smooth seasonal risk premium is able to explain the price movement to better extend than constant or zero risk premiums in the short end.

RMSE	θ	$\hat{\theta}^{OLS}$	$\hat{\theta}_t^{smooth}$	MAPE	θ	$\hat{\theta}^{OLS}$	$\hat{\theta}_t^{smooth}$
mean(θ)	0	0.042	0.039	mean(θ)	0	0.042	0.039
$\Delta(U_t, F_t)$	6.906	6.906	6.906	$\Delta(U_t, F_t)$	20.011	20.011	20.011
$\Delta(U_t, \hat{F}_t)$	13.979	12.495	7.388	$\Delta(U_t, \hat{F}_t)$	43.341	43.850	22.607
$\Delta(F, \hat{F}_t)$	9.975	9.293	3.745	$\Delta(F, \hat{F}_t)$	36.874	39.663	19.992

Table 7: Out-of-sample backtesting results for weekly WPF prices. Left: RMSE. Right: MAPE

RMSE	θ	$\hat{\theta}^{OLS}$	$\hat{\theta}_t^{smooth}$	MAPE	θ	$\hat{\theta}^{OLS}$	$\hat{\theta}_t^{smooth}$
mean(θ)	0	0.014	0.013	mean(θ)	0	0.014	0.013
$\Delta(U_t, F_t)$	2.447	2.447	2.447	$\Delta(U_t, F_t)$	12.479	12.479	12.479
$\Delta(U_t, \hat{F}_t)$	4.145	4.145	4.066	$\Delta(U_t, \hat{F}_t)$	20.091	17.323	17.245
$\Delta(F, \hat{F}_t)$	3.100	2.870	2.372	$\Delta(F, \hat{F}_t)$	13.805	12.645	8.460

Table 8: Out-of-sample backtesting results for monthly WPF prices. Left: RMSE. Right: MAPE

For out-of-sample backtesting results of monthly WPF contracts, presented in Table 8, the smooth MPR outperforms burn analysis and the OLS estimate using both criteria. Comparing with the actual average production as price estimate, smooth MPR delivers better results.

To test the sensitivity towards outliers we estimate cross-validated OLS and smooth MPR. The results can be taken from Table 9. The smooth MPR performs best for weekly contracts and RMSE, while the smooth MPR performs best for monthly contracts and MAPE, but on average similar or slightly worse than the average utilisation as price indicator.

RMSE	weekly WPF		monthly WPF	
	$\hat{\theta}_{CV}^{OLS}$	$\hat{\theta}_{t,CV}^{smooth}$	$\hat{\theta}_{CV}^{OLS}$	$\hat{\theta}_{t,CV}^{smooth}$
mean(θ)	0.042	0.040	0.014	0.013
$\Delta(U_t, F_t)$	6.906	6.906	2.447	2.447
$\Delta(U_t, \hat{F}_t)$	12.495	7.759	4.145	4.074
$\Delta(F, \hat{F}_t)$	9.293	4.085	2.870	2.496

MAPE	weekly WPF		monthly WPF	
	$\hat{\theta}_{CV}^{OLS}$	$\hat{\theta}_{t,CV}^{smooth}$	$\hat{\theta}_{CV}^{OLS}$	$\hat{\theta}_{t,CV}^{smooth}$
mean(θ)	0.042	0.040	0.014	0.013
$\Delta(U_t, F_t)$	20.011	20.011	12.479	12.479
$\Delta(U_t, \hat{F}_t)$	43.850	23.592	17.323	17.407
$\Delta(F, \hat{F}_t)$	39.663	20.556	12.645	9.760

Table 9: Out-of-sample cross-validated backtesting results for weekly and monthly WPF prices. Left: RMSE. Right: MAPE

5 Conclusion and further research

This applied research project presents a modern method to model weather derivatives with highly skewed underlying assets. We transform the underlying using logit-normal transformation and decompose the process into seasonality and a stationary mean-reverting process. To achieve overall best performance regarding normality, we propose a robust and smooth estimate of seasonal volatility, the smooth IQR. Further we compare our method to a stationary CARMA-Lévy model that incorporates the skewness of the underlying within jump-increments following a Variance Gamma distribution. We show that our method is preferred applying the Akaike information criterion for transformed models.

Further we calibrate the market price of risk (MPR) and find a MPR-smile, having positive MPR in im times of high seasonal volatility and negative MPR in times of low seasonal volatility. Employing a time varying smooth MPR estimate, our model outperforms burn analysis and OLS MPR estimates in out-of-sample analysis using mean absolute percentage error and root mean squared error as performance criteria. Results change only slightly after cross-validation.

Overall we show that our simple framework, employing data transformation and seasonal variance, allows for Gaussian risk drivers and captures the essential behaviour of WPFs to mimic the seasonal structure of MPR in a continuous time autoregressive model.

For further understanding of the price distribution, an analysis of tail behaviour could be done in future research. Adding another the meteorological variable wind speed as additional component into the model, might increase the forecasting performance. Another aspect for further research is location specific pricing. To enhance more precise hedging, location specific prices and measurements relative to the national WPF price and index are required. That would increase the trust in the product and its value for hedging purposes and hence, increase the trading volume and market liquidity.

References

- Aigner, D. J., Amemiya, T. & Poirier, D. J. (1976), ‘On the estimation of production frontiers: Maximum likelihood estimation of the parameters of a discontinuous density function’, *International Economic Review* **17**(2), 377–396.
- Akaike, H. (1978), ‘On the likelihood of a time series model’, *Journal of the Royal Statistical Society* **27**(D), 217–235.
- Alexandridis, A. & Zapranis, A. (2013), ‘Wind derivatives: Modeling and pricing’, *Computational Economics* **41**(3), 299–326.
- Barndorff-Nielsen, O. (1978), ‘Hyperbolic distributions and distributions on hyperbolae’, *Scandinavian Journal of Statistics* **5**(3), 151–157.
- Barndorff-Nielsen, O. E., Benth, F. E. & Veraart, A. E. D. (2013), ‘Modelling energy spot prices by volatility modulated lévy-driven volterra processes’, *Bernoulli* **19**(3), 803–845.
- Benth, F. E., Ebbeler, S. & Kiesel, R. (2014), ‘Indifference pricing of weather futures based on electricity futures’.
- Benth, F. E., Härdle, W. K. & López-Cabrera, B. (2011), *Pricing of Asian temperature risk*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 163–199.
- Benth, F. E., Klüppelberg, C., Müller, G. & Vos, L. (2014), ‘Futures pricing in electricity markets based on stable carma spot models’, *Energy Economics* **44**, 392–406.
- Benth, F. E. & Ortiz-Latorre, S. (2014), ‘A pricing measure to explain the risk premium in power markets’, *SIAM Journal on Financial Mathematics* **5**(1), 685–728.
- Benth, F. E. & Šaltytė Benth, J. (2009), ‘Dynamic pricing of wind futures’, *Energy Economics* **31**(1), 16 – 24.
- Benth, F. E., Šaltytė Benth, J. & Koekebacker, S. (2007), ‘Putting a price on temperature’, *Scandinavian Journal of Statistics* **34**(4), 746–767.
- Bowman, A. & Azzalini, A. (1997), *Applied smoothing techniques for data analysis : The kernel approach with s-plus illustrations*, 1 edn, Oxford: Clarendon Press.
- Breckling, J. & Chambers, R. (1988), ‘M-quantiles’, *Biometrika* **75**(4), 761.

- Brockwell, P. J. & Hannig, J. (2010), ‘Carma(p,q) generalized random processes’, *Journal of Statistical Planning and Inference* **140**(12), 3613 – 3618.
- Campbell, S. D. & Diebold, F. X. (2005), ‘Weather forecasting for weather derivatives’, *Journal of the American Statistical Association* **100**(469), 6–16.
- Gersema, G. & Wozabal, D. (2017), ‘An equilibrium pricing model for wind power futures’, *Energy Economics* **65**, 64–74.
- Groll, A., López-Cabrera, B. & Meyer-Brandis, T. (2016), ‘A consistent two-factor model for pricing temperature derivatives’, *Energy Economics* **55**, 112–126.
- Härdle, W. K. & López-Cabrera, B. (2012), ‘The implied market price of weather risk’, *Applied Mathematical Finance* **19**(1), 59–95.
- Härdle, W. K., López-Cabrera, B., Okhrin, O. & Wang, W. (2016), ‘Localizing temperature risk’, *Journal of the American Statistical Association* **111**(516), 1491–1508.
- Härdle, W. & Osipenko, M. (2017), Dynamic valuation of weather derivatives under default risk, SFB 649 Discussion Papers SFB649DP2017-005, Humboldt University, Collaborative Research Center 649.
- Jewson, S. & Brix, A. (2005), *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations*, Cambridge University Press.
- Karatzas, I. & Shreve, S. E. (1991), *Brownian motion and stochastic calculus*, 2 edn, Springer: New York.
- López-Cabrera, B., Odening, M. & Ritter, M. (2013), ‘Pricing rainfall futures at the cme’, *Journal of Banking & Finance* **37**(11), 4286–4298.
- Newey, W. K. & Powell, J. L. (1987), ‘Asymmetric least squares estimation and testing’, *Econometrica* **55**(4), 819–847.
- Pérez-González, F. & Yun, H. (2013), ‘Risk management and firm value: Evidence from weather derivatives’, *The Journal of Finance* **68**(5), 2143–2176.
- Pinson, P. (2012), ‘Very-short-term probabilistic forecasting of wind power with generalized logit–normal distributions’, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **61**(4), 555–576.

- Ritter, M., Shen, Z., López-Cabrera, B., Odening, M. & Deckert, L. (2015*a*), ‘Designing an index for assessing wind energy potential’, *Renewable Energy* **83**, 416 – 424.
- Ritter, M., Shen, Z., López-Cabrera, B., Odening, M. & Deckert, L. (2015*b*), ‘A new approach to assess wind energy potential’, *Energy Procedia* **75**, 671 – 676. Clean, Efficient and Affordable Energy for a Sustainable Future: The 7th International Conference on Applied Energy (ICAE2015).
- Schlemm, E. & Stelzer, R. (2012), ‘Quasi maximum likelihood estimation for strongly mixing state space models and multivariate lévy-driven carma processes’, *Electron. J. Statist.* **6**, 2185–2234.
- Shiryaev, A. N. (2001), *Essentials of Stochastic Finance - Facts, Models, Theory*, 3 edn, World Scientific, Singapore.
- Veraart, A. E. D. (2016), *Modelling the Impact of Wind Power Production on Electricity Prices by Regime-Switching Lévy Semistationary Processes*, Springer International Publishing, Cham, pp. 321–340.
- Wahba, G. (1990), *4. Estimating the Smoothing Parameter*, CBMS-NSF Regional Conference Series in Applied Mathematics, pp. 45–65.
- Weron, R. (2008), ‘Market price of risk implied by asian-style electricity options and futures’, *Energy Economics* **30**(3), 1098 – 1115.
- Ziel, F., Croonenbroeck, C. & Ambach, D. (2016), ‘Forecasting wind power—modeling periodic and non-linear effects under conditional heteroscedasticity’, *Applied Energy* **177**, 285–297.

SFB 649 Discussion Paper Series 2017

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Fake Alpha" by Marcel Müller, Tobias Rosenberger and Marliese Uhrig-Homburg, January 2017.
- 002 "Estimating location values of agricultural land" by Georg Helbing, Zhiwei Shen, Martin Odening and Matthias Ritter, January 2017.
- 003 "FRM: a Financial Risk Meter based on penalizing tail events occurrence" by Lining Yu, Wolfgang Karl Härdle, Lukas Borke and Thijs Benschop, January 2017.
- 004 "Tail event driven networks of SIFIs" by Cathy Yi-Hsuan Chen, Wolfgang Karl Härdle and Yarema Okhrin, January 2017.
- 005 "Dynamic Valuation of Weather Derivatives under Default Risk" by Wolfgang Karl Härdle and Maria Osipenko, February 2017.
- 006 "RiskAnalytics: an R package for real time processing of Nasdaq and Yahoo finance data and parallelized quantile lasso regression methods" by Lukas Borke, February 2017.
- 007 "Testing Missing at Random using Instrumental Variables" by Christoph Breunig, February 2017.
- 008 "GitHub API based QuantNet Mining infrastructure in R" by Lukas Borke and Wolfgang K. Härdle, February 2017.
- 009 "The Economics of German Unification after Twenty-five Years: Lessons for Korea" by Michael C. Burda and Mark Weder, April 2017.
- 010 "Data Science & Digital Society" by Cathy Yi-Hsuan Chen and Wolfgang Karl Härdle, May 2017.
- 011 "The impact of news on US household inflation expectations" by Shih-Kang Chao, Wolfgang Karl Härdle, Jeffrey Sheen, Stefan Trück and Ben Zhe Wang, May 2017.
- 012 "Industry Interdependency Dynamics in a Network Context" by Ya Qian, Wolfgang Karl Härdle and Cathy Yi-Hsuan Chen, May 2017.
- 013 "Adaptive weights clustering of research papers" by Larisa Adamyan, Kirill Efimov, Cathy Yi-Hsuan Chen, Wolfgang K. Härdle, July 2017.
- 014 "Investing with cryptocurrencies - A liquidity constrained investment approach" by Simon Trimborn, Mingyang Li and Wolfgang Karl Härdle, July 2017.
- 015 "(Un)expected Monetary Policy Shocks and Term Premia" by Martin Kliem and Alexander Meyer-Gohde, July 2017.
- 016 "Conditional moment restrictions and the role of density information in estimated structural models" by Andreas Tryphonides, July 2017.
- 017 "Generalized Entropy and Model Uncertainty" by Alexander Meyer-Gohde, August 2017.
- 018 "Social Security Contributions and the Business Cycle" by Anna Almosova, Michael C. Burda and Simon Voigts, August 2017.
- 019 "Racial/Ethnic Differences In Non-Work At Work" by Daniel S. Hamermesh, Katie R. Genadek and Michael C. Burda, August 2017.
- 020 "Pricing Green Financial Products" by Awdesch Melzer, Wolfgang K. Härdle and Brenda López Cabrera, August 2017.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

